

TESTING THE BINARY HYPOTHESIS: PULSAR TIMING CONSTRAINTS ON SUPERMASSIVE BLACK HOLE BINARY CANDIDATES

ALBERTO SESANA¹, ZOLTÁN HAIMAN^{2,3}, BENCE KOCSIS⁴ & LUKE ZOLTAN KELLEY⁵

¹School of Physics and Astronomy and Institute of Gravitational Wave Astronomy, University of Birmingham, Edgbaston B15 2TT, United Kingdom

²Department of Astronomy, Columbia University, 550 West 120th Street, New York, NY 10027, USA

³Department of Physics, New York University, New York, NY 10003, USA

⁴Institute of Physics, Eötvös University, Pázmány P. s. 1/A, Budapest, 1117, Hungary and

⁵Harvard University, Center for Astrophysics, Cambridge, MA 02138, USA

Draft version April 6, 2017

ABSTRACT

The advent of time domain astronomy is revolutionizing our understanding of the Universe. Programs such as the Catalina Real-time Transient Survey (CRTS) or the Palomar Transient Factory (PTF) surveyed millions of objects for several years, allowing variability studies on large statistical samples. The inspection of $\approx 250k$ quasars in CRTS resulted in a catalogue of 111 potentially periodic sources, put forward as supermassive black hole binary (SMBHB) candidates. A similar investigation on PTF data yielded 33 candidates from a sample of 33k objects. Working under the SMBHB hypothesis, we compute the implied SMBHB merger rate and we use it to construct the expected gravitational wave background (GWB) at nano-Hz frequencies, probed by pulsar timing arrays (PTAs). After correcting for incompleteness, we find that the implied GWB exceeds the current most stringent PTA upper limits by almost an order of magnitude for the CRTS sample, and by at least a factor of 2-3 for the PTF sample. Eccentricity and/or coupling with the environment can only suppress the signal by a factor of ≈ 2 in the relevant frequency range and cannot account for the inconsistency. The inferred GWB can be reconciled with the PTA upper limits only if nearly all 111 candidates are false positives and AGN hosting a single SMBH exhibit prominent variability without the need of an SMBH companion. Alternatively, alleviating this severe tension would require that the typical black hole mass has been overestimated by a factor of $\gtrsim 4$, very low typical mass ratios $q \equiv M_2/M_1 \lesssim 0.01$, or else that the loudest gravitational wave sources are preferentially false positives.

Keywords: black hole physics – gravitational waves – quasars: general – surveys

1. INTRODUCTION

In the past two decades, it has been realised that supermassive black holes (SMBHs) are a fundamental ingredient of galaxy formation and evolution (see, e.g., [Kauffmann & Haehnelt 2000](#); [Croton et al. 2006](#)), and it is now well established that possibly all massive galaxies host an SMBH at their centre (see [Kormendy & Ho 2013](#), and references therein). In the standard Λ CDM cosmology, structure formation proceeds in a hierarchical fashion, whereby galaxies frequently merge with each other, progressively growing their mass ([White & Rees 1978](#)). Following the merger of two galaxies, the SMBH hosted in their nuclei sink to the centre of the merger remnant because of dynamical friction (DF), eventually forming a SMBH binary (SMBHB).

The evolution of SMBHBs was first sketched out by [Begelman et al. \(1980\)](#). After the initial DF phase, the two SMBHBs become bound at parsec scales forming a Keplerian system. At this point DF ceases to be efficient and in the absence of any other physical mechanism at play, the binary would stall. Because the average massive galaxy suffers more than one major merger in its assembly, in this scenario virtually all galaxies would host parsec scale SMBHBs. Galactic nuclei, are however densely populated with stars and also contain gas. It has been shown that both three-body ejection of ambient stars ([Berczik et al. 2006](#); [Khan et al. 2011](#); [Preto et al. 2011](#); [Vasiliev et al. 2015](#); [Sesana & Khan 2015](#)) and interaction with gaseous clumps ([Goicovic et al. 2016](#)) or with a massive circumbinary disc ([Escala et al. 2005](#); [MacFadyen & Milosavljević 2008](#); [Hayasaki 2009](#); [Cuadra et al. 2009](#); [Roedig et al. 2011](#); [Nixon et al. 2011](#); [Roedig et al. 2012](#);

[Shi et al. 2012](#); [Farris et al. 2014](#); [Miranda et al. 2017a](#); [Tang et al. 2017](#); see also the reviews by [Dotti et al. 2012](#); [Mayer 2013](#)) may provide efficient ways to shrink the binary down to centi-parsec scales, where efficient gravitational wave (GW) emission takes over, leading to final coalescence. Still, binary hardening time-scales can be as long as Gyrs ([Sesana & Khan 2015](#); [Vasiliev et al. 2015](#)), implying a substantial population of sub-parsec SMBHB lurking in galactic nuclei.

Sub-pc SMBHBs are extremely elusive objects (see [Dotti et al. 2012](#); [Komossa & Zensus 2016](#), for recent reviews). At extragalactic distances, their angular size is well below the resolution of any current instrument, making direct imaging impossible. Conversely, there is an increasing number of detections of SMBH *pairs* (i.e. SMBHBs still not gravitationally bound to each others) at hundred pc to kpc separations in merging galaxies, which are their natural progenitors (e.g. [Hennawi et al. 2010](#); [Comerford et al. 2013](#)). With direct imaging impractical, other avenues to observe sub-pc SMBHBs have been pursued, namely spectroscopic identification and time variability. Because of the high (> 1000 km/s) typical orbital velocities, SMBHBs have tentatively been identified with systems showing significant offsets of the broad line emission lines compared to the reference provided by narrow emission lines, and/or with frequency shifts of the broad lines over time ([Tsalmantza et al. 2011](#); [Eracleous et al. 2012](#); [Decarli et al. 2013](#); [Ju et al. 2013](#); [Shen et al. 2013](#); [Wang et al. 2017](#); [Runnoe et al. 2017](#)). The latter, in fact, are generated within the host galaxies at hundred of parsecs from the central SMBHBs, whereas the former are generated by gas bound to the SMBHBs. If the SMBHBs have a significant velocity com-

pared to the galaxy rest frame, and the broad emission region is bound to the individual SMBHBs, then broad lines will have an extra redshift/blueshift compared to their narrow counterparts. Note, however, that for very compact SMBHBs those broad lines might in fact be generated within the circumbinary disk (Lu et al. 2016), questioning this interpretation.

With the advent of time domain astronomy, the identification of SMBHBs via periodic variability has been put forward (Haiman et al. 2009). The rationale for selecting candidates in this way is that if gas is being accreted onto a SMBHB, the orbital period of the system could translate into periodic variability of the emitted luminosity. In fact, detailed hydrodynamical simulations show that SMBHBs embedded in circumbinary discs carve a cavity in the gas distribution and the gas streams from the cavity edge onto the binary in a periodic fashion (e.g. Artymowicz & Lubow 1996; MacFadyen & Milosavljević 2008; Hayasaki et al. 2007; Cuadra et al. 2009; Roedig et al. 2011; Shi et al. 2012; Noble et al. 2012; D’Orazio et al. 2013; Farris et al. 2014; Shi & Krolik 2015).

Whether such periodic streaming translates into periodicity in the emitted luminosity is less clear. Moreover, it has been pointed out that such periodicity would mostly impact the UV and X-ray emission from the binary, whereas the optical emission coming from the circumbinary disc might be relatively steady (Sesana et al. 2012; Farris et al. 2015), except for the most massive ($M \gtrsim 10^9 M_\odot$) SMBHBs for which the optical emission arises from gas bound to the individual black holes (D’Orazio et al. 2015b). Despite these uncertainties, active galactic nuclei (AGN) showing periodic variability are viable candidates for hosting SMBHBs. Based on this hypothesis, Graham et al. (2015) (hereafter G15) proposed 111 SMBHB candidates by inspecting the light curves of 243,486 quasars identified in the Catalina Real-time Transient Survey (CRTS, Drake et al. 2009). In a similar effort, Charisi et al. (2016) (hereafter C16) identified 33 SMBHB candidates, among 35,383 spectroscopically confirmed quasars in the Palomar Transient Factory (PTF; Rau et al. 2009) survey, with somewhat lower magnitudes and shorter periods than G15. Both groups presented a thorough analysis demonstrating a large excess of periodic sources at odds with the expectations of standard AGN-variability models.

The statistical significance of the detected periodicity depends strongly on the stochastic noise model for underlying quasar variability. For example, Vaughan et al. (2016) have recently showed that red noise models can naturally lead to frequent false positives in periodicity searches, especially at inferred periods comparable to the length of the data stream. At the same time, observations are expected to yield binaries preferentially at lower frequencies where they spend the largest fraction of their lifetimes. On the other hand, pure red noise (i.e. a single f^{-2} power-law noise power spectrum) is a poor description of quasar variability in general. A better fit to the power spectra observed for large quasar samples is the damped-random walk (DRW) noise model, in which the red noise power spectrum flattens to white noise ($\propto f^0$) at low frequencies (e.g. MacLeod et al. 2010). Adopting a DRW reduces the stochastic noise power at low frequencies, and therefore translate to a higher significance of an observed periodicity. As a result, the significance of the inferred periodicities depend strongly on the poorly constrained underlying noise model (see further discussion of this point in, e.g. Kelly et al. 2014; Charisi et al. 2016; Vaughan et al. 2016; Kozłowski 2017).

In this paper we test the SMBHB hypothesis *on physical*

grounds. SMBHBs are powerful emitters of nHz gravitational waves (GWs), which are currently probed by pulsar timing arrays (PTAs Foster & Backer 1990). There are three ‘regional’ PTAs currently in operation: the European Pulsar Timing Array (EPTA, Desvignes et al. 2016), the Parkes Pulsar Timing Array (PPTA, Reardon et al. 2016) and the North American Nanohertz Observatory for Gravitational Waves (NANOGrav, Arzoumanian et al. 2015). These three collaborations share data under the aegis of the International Pulsar Timing Array (IPTA, Verbiest et al. 2016), with the goal of improving their combined sensitivity. The collective signal from a cosmic population of SMBHBs results in a stochastic GW background (GWB, Rajagopal & Romani 1995; Jaffe & Backer 2003; Wyithe & Loeb 2003; Sesana et al. 2008), but particularly massive/nearby systems might emit loud signals individually resolvable above the median background level (Sesana et al. 2009).

Recent PTA efforts resulted in several upper limits both for a GWB (Shannon et al. 2015; Lentati et al. 2015; Arzoumanian et al. 2016) and for individual sources (Arzoumanian et al. 2014; Zhu et al. 2014; Babak et al. 2016). Both G15 and C16 showed that each SMBHB candidate identified in CRTS is individually compatible with current PTA upper limits. This is partly because residuals from those individual SMBHBs are simply too small and partly because their GW emission lies at frequencies above $> 10^{-8}$ Hz, which is above the PTA detection ‘sweet spot’ at $\approx 10^{-8.3}$ Hz. The 10^{-8} Hz lower limit to the GW frequency, however, is only a selection effect due to the length of the CRTS data stream (9 yrs). We show that, when properly converted into a SMBHB merger rate and extrapolated to lower frequencies, both samples are severely inconsistent with PTA upper limits. The inconsistency becomes naturally worse when the computation is corrected for incompleteness of the G15 sample. This implies that either the masses of the SMBHBs have been severely overestimated or most of the candidates are indeed not binaries with orbital periods matching their lightcurve variability, which therefore has a different physical origin. We show that our conclusions are not severely affected by physical processes potentially capable of suppressing the low frequency GW signal, such as significant eccentricities, or strong environmental coupling. This paper is organized as follows. We concentrate our investigation on the G15 SMBHB candidate sample, which is presented in § 2 and used in § 3 to reconstruct the merger rate of SMBHBs throughout cosmic history. The implied GWB is computed in § 4 and compared to current PTA upper limits. In § 5 we apply the same formalism to the C16 sample. The main results are summarized in § 6. Throughout this paper, we adopt a concordance Λ CDM cosmology with $h = 0.679$, $\Omega_M = 0.306$, $\Omega_\Lambda = 0.694$ (Planck Collaboration XIII 2016).

2. THE CATALINA SURVEY SMBHB CANDIDATE SAMPLE

CRTS is a time-domain survey periodically scanning 33,000 deg² (80% of the whole sky). The data release used by G15 contains the light curves of millions of objects monitored over nine years. Objects have been cross-checked with the 1M quasar catalogue¹ to identify more than 300k spectroscopically confirmed quasars, of which 243,486 have sufficient light curve coverage for a periodicity search. Among these sources, 111 have been flagged as periodically varying

¹ <http://quasars.org/milliquas.htm>

and have been proposed as potential SMBHB candidates. For most of these systems, the values of the estimated SMBHB mass, redshift and orbital period is provided by G15 in tabular form.

2.1. Observational properties

Of the 111 candidates presented by G15, we consider only the 98 with a reported measurement of the mass. In the following we will conservatively assume that this is the total SMBHB mass ($M_t = M_1 + M_2$). The mass and redshift distributions of these 98 systems are plotted in Fig. 1. Interestingly, the number of sources peaks at $z \approx 1$, rapidly declining to zero at $z > 2$. In the binary hypothesis, this can be explained by a selection effect. G15's search is sensitive to observed periods between 20-300 weeks and has a fixed magnitude limit. As a result, at higher redshifts, they could only find binaries with shorter rest-frame orbital periods (< 2 yr at $z > 2$), and with increasingly large masses. These massive, compact binaries have decoupled from their circumbinary discs, and are likely well inside the GW-driven inspiral regime (Haïman et al. 2009), with very short ($\lesssim 10^4$ yr) inspiral times. As a result, they will be exceedingly rare. Additionally, binaries can take several Gyrs to overcome the ‘final parsec problem’, so that the distribution of SMBHBs might peak at lower redshift with respect to the quasar luminosity function.

Irrespective of the reason for the low- z bias in the G15 sample, previous studies have found that the overwhelming majority of the GW signal comes from low-redshift systems with masses larger than $10^8 M_\odot$ (see, e.g. Sesana et al. 2008; Kocsis & Sesana 2011; Ravi et al. 2015; Simon & Burke-Spolaor 2016). Following Sesana (2013b) we will therefore first consider only sources with $z < 1.3$; the G15 sample should not be severely incomplete at these low redshifts. The SMBHB candidates are extremely massive, and the peak of the mass distribution is at $M_t \approx 10^9 M_\odot$ with a long tail extending to $10^{10} M_\odot$. Note that, at a given frequency, more massive binaries coalesce faster; this means that, when converted into a merger rate, the distribution shown in the lower panel of Fig. 1 will be skewed to even higher masses. Although the resulting skewness is likely due to incompleteness of the CRTS survey at low masses (see further discussion in Section 3.1), we will see that the implied merger rate has critical consequences for our study. Along the same lines, we also notice that the G15 sample features several massive binaries at fairly high redshift. At equal observed frequency, binaries at higher redshift are emitting at *higher* rest frame frequency, meaning that the rest frame coalescence times to which CRTS has access are *shorter*. This means that they potentially contribute significantly to the inferred cosmic merger rate and to the GW signal. We will indeed show below that this is the case, contrary to what was found in the aforementioned theoretical studies.

2.2. Assigning individual masses

The CRTS catalogue only provides an estimate of the total SMBHB mass. To compute the associated GW signal the mass ratio $q = M_2/M_1$ (where by definition $M_1 > M_2$) of the sources is also needed. Inferring the q distribution of observable SMBHBs is far from being a trivial task, and we will see that it is an important factor in assessing PTA constraints. We describe in the following some of the subtleties that come into play, which led us to consider three different scenarios.

In comparing their sample with PTA upper limits, G15 assumed a typical $q = 1/2$. Although this might appear a reasonable choice, galaxy formation models predict indeed a

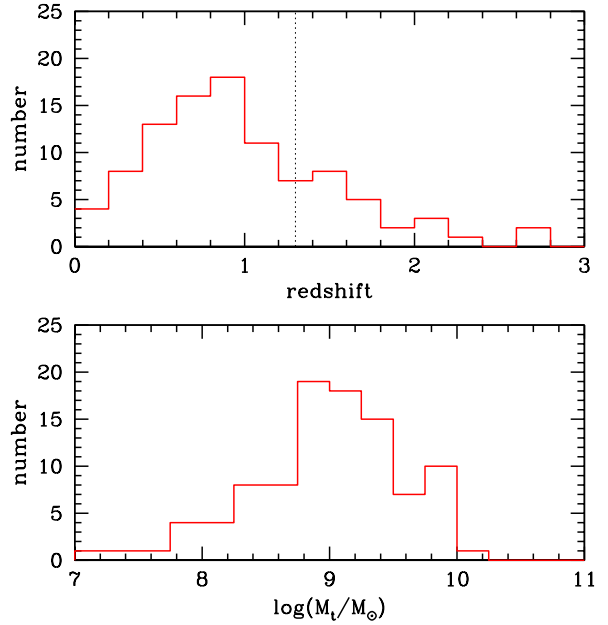


Figure 1. Redshift (upper panel) and total mass (lower panel) distribution of the SMBHB candidates identified by Graham et al. (2015) in the CRTS. Out of the 111 candidates, only the 98 with reported black hole mass estimates are shown. The dotted vertical line in the upper panel at $z = 1.3$ marks the higher redshift considered in SMBHB population models by Sesana (2013b).

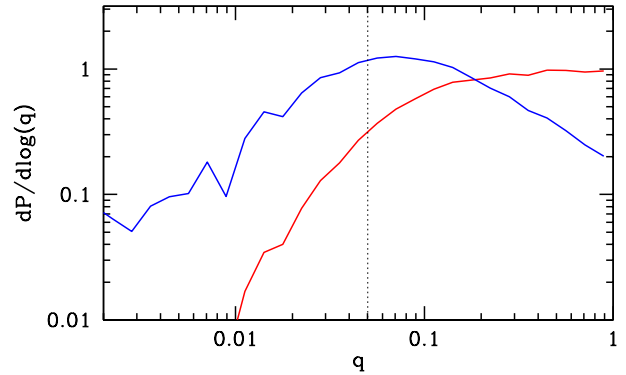


Figure 2. Probability distribution of the mass ratio of merging SMBHBs in the Millennium Run considering only systems with $M_t > 3 \times 10^8 M_\odot$ merging at $z < 2$ (red line) compared to the q probability distribution of a SMBHB observed in a selected frequency range (blue line). The vertical dotted line marks $q = 0.05$, above which accretion induced periodicity is seen in numerical simulations of SMBHBs in circumbinary disks (see text for details).

larger range of q , as shown in Fig. 2. Here we consider all the N_m SMBHBs with $M_t > 3 \times 10^8 M_\odot$ merging at $z < 2$ in a modified version of the galaxy formation models of Guo et al. (2013) implemented on the Millennium Simulation (Springel et al. 2005). Modifications included populating merging galaxies with SMBHBs according to a specific scaling relation (in this case the M – σ relation from Gültekin et al. 2009), and adding appropriate delays between galaxy mergers and SMBHB mergers to account for dynamical friction and SMBHB hardening. Plotted is the probability distribution $\mathcal{P}(\log q)$ (normalized so that $\int \mathcal{P}(\log q) d\log q = 1$) of those merging systems. The distribution is essentially flat down to $q = 0.1$, and drops at lower mass ratios. We see however that mergers down to $q \approx 0.01$ are still possible. Similar distributions, albeit sometimes flatter and extending to even lower q , have been found in semi-analytic models (e.g. Barausse 2012) and are produced by full cosmological hydrodynamical simu-

lations like Illustris (Kelley et al. 2017).

If we take a snapshot of the sky, the q distribution of SMBHB observed at a given frequency is not the same as that of the merging systems. In fact, if we assume for simplicity GW driven, circular binaries, the residence time at a given frequency is proportional to (see, e.g. Sesana et al. 2008)

$$\frac{dt}{df} = \frac{5}{96\pi^{8/3}} \left(\frac{GM}{c^3} \right)^{-5/3} f^{-11/3}, \quad (1)$$

where $\mathcal{M} = M_1^{3/5} M_2^{3/5} / M_t^{1/5} = M_1 q^{3/5} / (1+q)^{1/5}$ is the binary chirp mass. Therefore, for a given M_t (and ignoring the $1+q \approx 1$ factor) $dt/df \propto q^{-1}$. Since the number of observable binaries in a given frequency window is $dN/df \propto N_m \times dt/df$, their actual q distribution is proportional to $\mathcal{P}(\log q)/q$. This is shown by the blue line in Fig. 2: the q distribution is now skewed towards lower values, peaking at $q \lesssim 0.1$.

Small q binaries should therefore be quite common. Whether they are observable as periodic quasars, however, is less clear. In the standard picture, periodicity in the light curve is associated with variability in the accretion onto a SMBHB as $L(t) \propto \dot{M}(t)$. The typical variability of 0.2–0.5 magnitudes observed in CRTS thus corresponds to accretion fluctuations (and relative luminosity fluxes) of about 20%–60%, necessitating a relatively high q . For example, Farris et al. (2014) find that $q > 0.05$ is needed to get a distinctive variability pattern in the accretion rate (which is marked in Fig. 2). On the other hand, D’Orazio et al. (2015b) proposed that the sinusoidal behaviour of the light curve is given by relativistic Doppler boosting. This model works only for $q \lesssim 0.1$, making it almost complementary to the accretion-induced variability. Besides the technical details of the source of variability, the large accretion rates implied by the CRTS sample of bright quasars are generally found to be triggered by major mergers (e.g. Croton et al. 2006). Therefore the q distribution of SMBHBs hosted in bright quasars might be biased-high.

In light of these uncertainties, we construct three mass ratio models:

- **Model 1:** q drawn from a log-normal distribution with $\sigma_{\log q} = 0.5\text{dex}$, peaked at $\log q = 0$ (and we consider only $\log q < 0$). This model is representative of a biased-high q distribution, and is similar to the $q = 1/2$ study case considered in G15;
- **Model 2:** q is drawn from the distribution plotted in Fig. 2 (blue curve), but with a minimum cut-off placed at $q = 0.05$. This model is representative of accretion induced periodicity;
- **Model 3:** q is drawn from the distribution plotted in Fig. 2 (blue curve), without cut-off. This model privileges binaries with $q \lesssim 0.1$ and is representative of the Doppler boosting scenario.

We also investigate the effect of the intrinsic uncertainty on the mass measurement given in G15, and for each of the models above we consider two scenarios. In scenario ‘a’ we consider the G15 masses to be exact; in scenario ‘b’ we assign to each mass measurement an error of $\sigma_{\log M_t} = 0.3\text{dex}$ (consistent with the mass-determination method used in G15; Shen et al. 2008). We therefore have six models in total, that we label Model 1a/b, Model 2a/b, Model 3a/b.

Note that Model 3a/b imply a moderate inclination with respect to the observer line of sight (otherwise the boosting would not be efficient enough, see details in D’Orazio

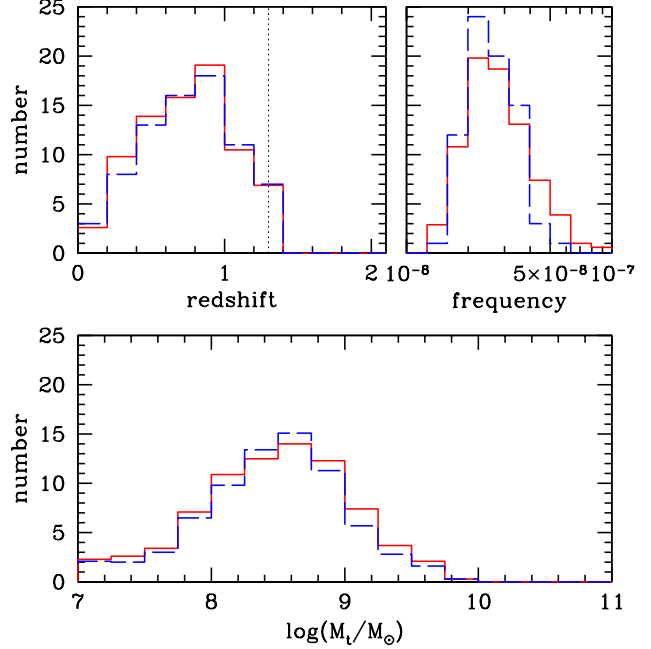


Figure 3. Comparison between the 98 SMBHB candidates in the CRTS, and a mock population of SMBHBs constructed from the merger rate implied by the CRTS systems. *Upper left:* redshift distribution. *Upper right:* rest-frame frequency distribution. *Lower panel:* chirp mass distribution. The chirp mass distribution is constructed assuming Model 1b (see text for details). In all panels, solid red histograms are the CRTS sample, dashed blue histograms are an average over 10 Monte Carlo realisations of Model 1b.

et al. 2015b). This implies that the CRTS observed sources would be an incomplete sample (including only systems with favourable inclination) of the underlying SMBHB population. We do not attempt to address this incompleteness in the following analysis, which makes our estimates for Model 3a/b conservative, and accounts for the possibility of accretion induced variability at low q .

3. BUILDING MOCK SMBHB POPULATIONS

For each of the six models enumerated in the previous section, we take the 98 CRTS candidates with estimated total mass and we assign them M_t and q by drawing from the respective distributions. We repeat the procedure 1000 times, to get a statistical ensemble of SMBHBs samples under each of the model prescriptions.

Each individual realisation of the 98 CRTS SMBHB candidates can be used to infer a SMBHB merger rate as follows. Neglecting for the moment completeness issues, the output of the CRTS can be treated as an all sky population of binaries with given masses and redshifts, emitting in some frequency range. We note that, because of the limited time span of the data, CRTS is sensitive only to SMBHBs with *observed* orbital periods shorter than the threshold value $\tilde{P}_{\text{orb}} = 6\text{yrs}$. This corresponds to a *rest frame* GW frequency $\tilde{f}_r = 2 \times \tilde{P}_{\text{orb}}^{-1} (1+z)$. The rest frame coalescence time-scale of a (circular, GW driven) SMBHB emitting at \tilde{f}_r is

$$T_c(\tilde{f}_r) = \frac{5}{256} \left(\frac{GM}{c^3} \right)^{-5/3} (\pi \tilde{f}_r)^{-8/3}. \quad (2)$$

Suppose now we have N identical (in mass and redshift) binaries emitting at frequencies $f_r > \tilde{f}_r$ identified in the CRTS sample. By virtue of the continuity equation for SMBHB evo-

lution, we can convert this number, into a merger rate using:

$$\dot{N} \equiv \frac{dN}{dt_r} = \frac{N(f_r > \tilde{f}_r)}{T_c(\tilde{f}_r)}, \quad (3)$$

where we stress once again that $T_c(\tilde{f}_r)$ is the coalescence time at the longest orbital period probed by CRTS (\tilde{P}_{orb}), and not at the period of the specific binary candidate. We can then generalize the argument to a distribution of SMBHBs with different masses and redshifts and numerically construct a binned distribution $\Delta^2 \dot{N} / (\Delta \mathcal{M} \Delta z)$ by summing up the sources in the CRTS sample as follows,

$$\frac{\Delta^2 \dot{N}}{\Delta \mathcal{M} \Delta z} = \sum_{i \in \Delta \mathcal{M}} \sum_{i \in \Delta z} \frac{1}{\Delta \mathcal{M} \Delta z T_{c,i}(\tilde{f}_r)}. \quad (4)$$

Here $i = 1, \dots, 98$ is an index identifying the SMBHB candidates, and $T_{c,i}(\tilde{f}_r)$ is the coalescence time-scale of the i -th binaries according to equation (2). Note that although equation (4) depends on the choice of the bins $\Delta \mathcal{M}$ and Δz , when computing the total merger rate via equation (6) or the GW signal via equation (7) below, we integrate over $d\mathcal{M}$ and dz , and the final results are bin-independent. To ease the notation, we switch to the continuum and use the differential form $\Delta^2 \dot{N} / (\Delta \mathcal{M} \Delta z) \rightarrow d^3 N / (d\mathcal{M} dz dt_r)$. We stress, however that all computations have been performed numerically on binned distributions and the robustness of the results have been checked against the choice of bin sizes.

To check whether our merger rate is consistent with the observed CRTS SMBHB population, we can now construct the expected $d^3 N / (d\mathcal{M} dz df_r)$ as

$$\frac{d^3 N}{d\mathcal{M} dz df_r} = \frac{d^3 N}{d\mathcal{M} dz dt_r} \frac{dt_r}{df_r}, \quad (5)$$

where dt_r/df_r is the same as equation (1) but evaluated in the source rest-frame. We can then perform a Monte Carlo sampling of the distribution given by equation (5) and compare it to the original CRTS SMBHB candidate ensemble. The comparison is given in Fig. 3, where we show the chirp mass, rest-frame frequency and redshift distribution averaged over 10 realisations of Model 1b. The inferred merger rate numerically constructed as explained above is perfectly consistent with the observed CRTS SMBHB population. Equation (4) therefore provides a reliable estimate of the SMBHB merger rate implied by the observed candidates. We caution that this might differ from the *intrinsic* SMBHB merger rate as, in practice, completeness may depend on M , q , f_r , which distorts the distribution as discussed below.

3.1. Coalescence rates

We can now compute the *observed* differential merger rate per unit chirp mass, by integrating equation (4) in a given redshift range,

$$\frac{d^2 N}{d\mathcal{M} dt} = \int_{z_{\min}}^{z_{\max}} \frac{d^3 N}{d\mathcal{M} dz dt_r} \frac{dt_r}{dt} dz. \quad (6)$$

Results are shown in Fig. 4, where the merger rates of Model 1b and Model 3a are integrated in the redshift range $0 < z < 1.3$ and are compared to theoretical estimates from the literature. In particular, we consider an updated version of the observation-based models of Sesana (2013b), and two models extracted from the Millennium simulation (Springel et al. 2005) and described in Sesana et al. (2009). We consider

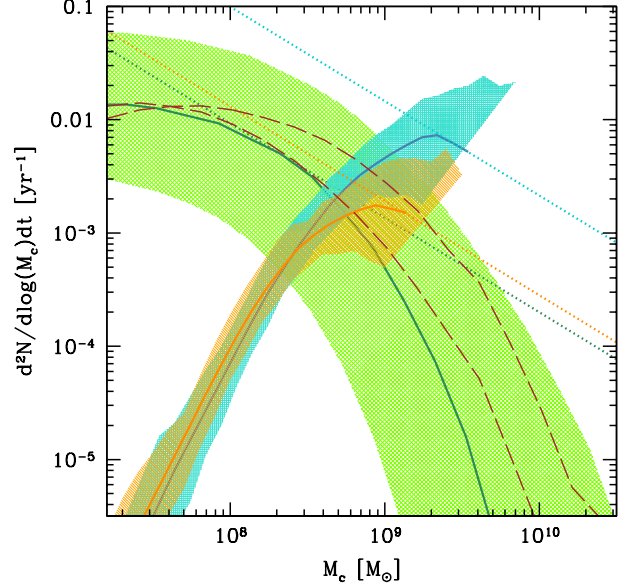


Figure 4. Merger rate mass function comparison. The green shaded area is the 95% interval produced by theoretical models presented in Sesana (2013b) and the solid dark green line is the median value. The blue and orange shaded areas are the 95% intervals based on the observed samples of periodic AGNs produced by Model 1b and Model 3a described in the text, and the thick blue and orange lines are the respective median values. The two dashed brown lines are two models selected from Sesana et al. (2009). The three dotted curves with a $-5/6$ slope are tangent to the median values predicted by the various models and show which chirp mass contributes the most to the GW signal (see text for details).

Model 1b and Model 3a because, among those we consider, they are extremal models in terms of SMBHB merger rate and GWB amplitude (as we will see later on). The comparison shows that Model 1b results in a merger rate that is hardly consistent with theoretical estimates. Although the *integrated* merger rate implied by the CRTS population is consistent with expectations from hierarchical clustering, the mass distributions do not match and are inconsistent at least at a 95% level at $\mathcal{M} > 2 \times 10^9 M_\odot$. Note that incompleteness of the sample can only increase the merger rate which makes this discrepancy worse at high masses; in fact, although undetected low mass systems might close the gap at $\mathcal{M} < 10^8 M_\odot$, any correction for incompleteness can only make the mismatch at the high mass end more severe (see discussion below). Conversely, Model 3a, in virtue of the typical small q , has a \mathcal{M} distribution that is still consistent with theoretical assembly models. In particular, we will see in equation (7) below that for a given merger rate, the gravitational wave strain is $h_c \propto \mathcal{M}^{5/6}$. This means that the largest contribution to the GW signal comes from the value of \mathcal{M} where the merger rate distribution is tangent to a line with $-5/6$ slope in the log-log plane, as depicted in Fig. 4. We see that the line tangent to the median of Model 3a is close to the one tangent to the median value of the theoretical models, implying a very similar GW strain (*cf.* left-most panel of Fig. 5). On the other hand, the inferred merger rate from the periodically variable AGN sample with Model 1b gives a systematically higher rate by an order of magnitude than the merger rate estimates from theoretical models, which results in a large over prediction of the GW signal (as later shown in Fig. 6). The sharp decline at $\mathcal{M} < 10^9 M_\odot$ might be due to CRTS incompleteness; lighter SMBHBs are intrinsically fainter and might be missing in the relatively shallow Catalina survey. We also stress that

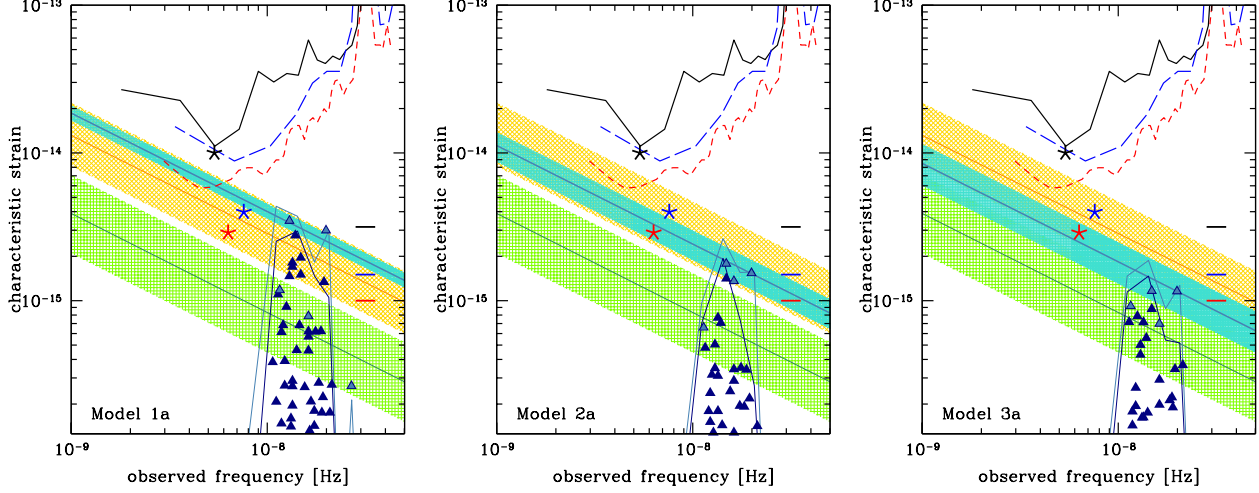


Figure 5. Characteristic amplitude of the GW signal for circular, GW-driven binaries. In each panel, the orange and green shaded areas are 68% intervals produced by selected models presented in [Sesana \(2013b\)](#) and [\(Sesana et al. 2016\)](#) (see text for details); the solid orange and green lines are their respective median values. The blue shaded area is the 68% interval produced by models based on the G15 SMBHB candidates and the solid blue line is the median value. The dark blue triangles illustrate one example of a Monte Carlo realisation of the individual GW signals of the G15 sample, according to the model labelled in each panel. The light blue triangles are sources that are individually louder than the contribution of all other binaries falling in their frequency bin (and therefore, potentially resolvable individually by a putative PTA with enough sensitivity), taken here to be $\Delta f = 3.17 \text{ nHz}$ (i.e. 10 yr^{-1}). The corresponding light blue line is the resulting overall signal, and the dark blue line is the level of the GWB after these putative resolvable sources are subtracted. The solid black, long-dashed blue and short-dashed red lines represent the sensitivities to a stochastic GW signal as a function of frequency for the EPTA, NANOGrav and PPTA respectively. The corresponding stars represent the 95% upper limit to the integrated GWB from each PTA, placed at the frequency at which it is most sensitive. The horizontal ticks are the extrapolation of those limits to a frequency of 1 yr^{-1} .

the merger rates inferred from CRTS have not been corrected for incompleteness in this comparison, therefore, the discrepancy is larger than shown in Fig. 4. We will discuss the effect of incompleteness on our results in § 4.3.

4. PTA IMPLICATIONS OF THE CATALINA SAMPLE

With a reliable estimate of the SMBHB merger rate in hand, we can now proceed to the computation of the expected GW signal. We start by considering circular, GW-driven binaries. Following [Phinney \(2001\)](#) and [Sesana et al. \(2008\)](#), we can write the overall GW signal as:

$$h_c^2(f) = \frac{4}{\pi f^2} \int_0^\infty dz \int_0^\infty d\mathcal{M} \frac{d^2 n}{dz d\mathcal{M}} \frac{1}{1+z} \frac{dE_{\text{gw}}(\mathcal{M})}{d \ln f_r}. \quad (7)$$

where the energy radiated per logarithmic frequency interval is

$$\frac{dE_{\text{gw}}}{d \ln f_r} = \frac{\pi^{2/3}}{3} \mathcal{M}^{5/3} f_r^{2/3} \quad (8)$$

and $d^2 n / dz d\mathcal{M}$ is the comoving number density of mergers per unit redshift and chirp mass and is related to the overall cosmic merger rate of equation (4) via

$$\frac{d^2 n}{dz d\mathcal{M}} = \frac{d^3 N}{d\mathcal{M} dz dt_r} \frac{dt_r}{dz} \frac{dz}{dV_c}. \quad (9)$$

The functions dt_r/dz and dz/dV_c are given by standard cosmology. We can therefore construct the expected GW signal by using equation (7) for all our models. For each model, we compute h_c from the 1000 realisations of the CRTS ensemble, to get a measurement of the uncertainty in the predicted signal amplitude.

Note that in the circular GW-driven scenario, the GW signal computed through equation (7) is a simple power-law with spectral slope $-2/3$ ([Phinney 2001](#)). The GW amplitude is thus customarily written as

$$h_c = A \left(\frac{f}{1 \text{ yr}^{-1}} \right)^{-2/3}, \quad (10)$$

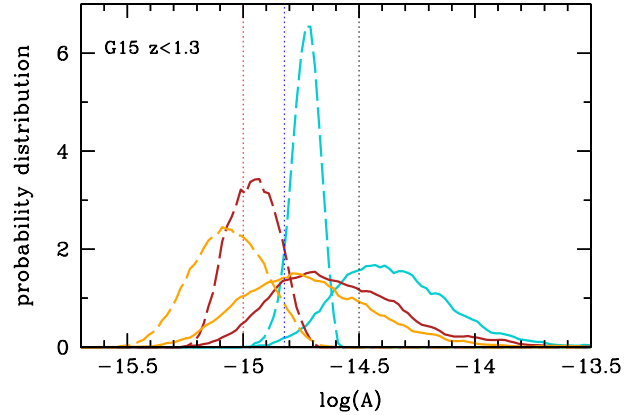


Figure 6. Distribution of the stochastic GWB characteristic amplitudes inferred from the CRTS sample under different assumptions about the masses and mass-ratios of the SMBHBs. The blue curves are distributions for Model 1a (dashed) and Model 1b (solid), brown curves are for Model 2a (dashed) and Model 2b (solid), and orange curves are for Model 3a (dashed) and Model 3b (solid). The vertical dotted lines are 95% upper limits given by EPTA (black), NANOGrav (blue) and PPTA (red).

and PTA results are quoted as limits on the amplitude normalization A at a nominal frequency of 1 yr^{-1} .

4.1. GW signal from SMBHB at $z < 1.3$

We first include in our calculation only SMBHBs emitting at $z < 1.3$. Results are shown in Fig. 5 for Model 1a, Model 2a and Model 3a, and are compared with the expected signal from observation based models featuring different SMBH-galaxy scaling relations. In particular, we pick from [Sesana \(2013b\)](#) a fairly ‘optimistic’ model in which the SMBH mass correlates with the host bulge mass following the relation proposed by [Kormendy & Ho \(2013\)](#). Note that the GW amplitude predicted by this model, $-15.1 < \log A < -14.7$ at 68% confidence, is already in tension with the current best 95% GWB upper limit at $\log A < -15$ ([Shannon et al.](#)

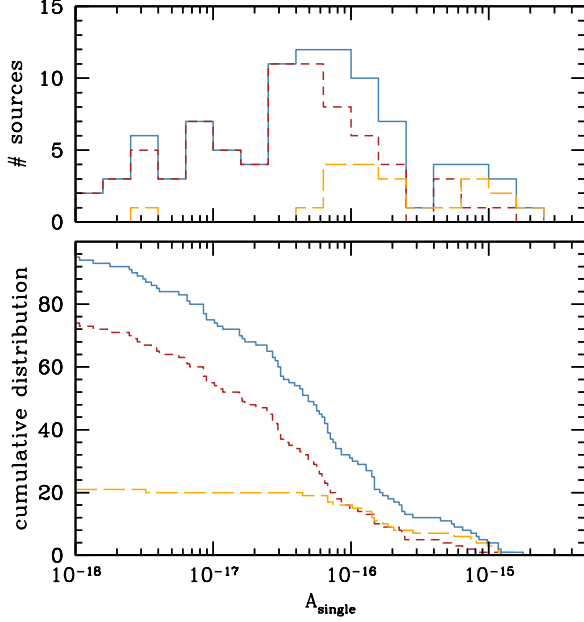


Figure 7. Contribution of each individual system in the G15 binary-candidate catalogue to the total GWB amplitude in Model 1b. The top panel shows the distribution of the median GWB contribution of each SMBHB (see text for description) for the overall sample (blue), for the $z < 1.3$ sub-sample (brown) and for the $z > 1.3$ sub-sample (orange). The lower panel shows, with the same line style, the cumulative distribution constructed from the histograms in the upper panel.

2015). We also consider an alternative model featuring a much more conservative SMBHB-host relation proposed by Shankar et al. (2016). This model predicts a GW amplitude of $-15.8 < \log A < -15.2$ at 68% confidence (Sesana et al. 2016).

Model 1a is inconsistent with current PTA upper limits, (Lentati et al. 2015; Shannon et al. 2015; Arzoumanian et al. 2016), in particular with both the PPTA (Shannon et al. 2015) and NANOGrav (Arzoumanian et al. 2016) results. In virtue of their progressively lower q distribution, Model 2a and 3a are still consistent with PTA measurements, although Model 2a only marginally.

Note, however, that these models assume no uncertainty in the mass measurement provided by G15. When this is included, the situation is quite different. The distribution of the GWB amplitude A at the fiducial frequency of 1 yr^{-1} , obtained from the 1000 Monte Carlo realisations of all models is shown in Fig. 6, and is compared to the current 95% PTA upper limits. Uncertainties in the mass measurement significantly broaden the A distribution, shifting $A_{50\%}$ (i.e. the median of the A distribution) to higher values. We obtain $\log(A_{50\%}) = -14.42, -14.65, -14.78$ for Model 1b, Model 2b and Model 3b respectively, all of which are inconsistent with PTA upper limits. We therefore conclude that, even when considering only events at $z < 1.3$ and ignoring any completeness issue, the list of SMBHB candidates proposed by G15 can hardly be reconciled with PTA upper limits.

4.2. Including SMBHBs at $z > 1.3$ and single source contribution

Even though the bulk of the GWB is expected to come from SMBHBs at $z < 1.5$ (Sesana et al. 2008; Ravi et al. 2015; Simon & Burke-Spolaor 2016), the G15 sample features several fairly massive candidates at higher redshifts. Note that, according to equation (7), the GW contribution depends on

Name	z	$P[\text{d}]$	$\log(\frac{M_t}{M_\odot})$	$A_s[10^{-15}]$
SDSS J164452.71+430752.2	1.715	2000	10.15	1.047
HS 0926+3608	2.150	1562	9.95	0.510
SDSS J140704.43+273556.6	2.222	1562	9.94	0.483
SDSS J131706.19+271416.7	2.672	1666	9.92	0.445
HS 1630+2355	0.821	2040	9.86	0.432
SDSS J134855.27-032141.4	2.099	1428	9.89	0.432
SDSS J114857.33+160023.1	1.224	1851	9.90	0.398
SDSS J092911.35+203708.5	1.845	1785	9.92	0.366
SDSS J160730.33+144904.3	1.800	1724	9.82	0.328
SDSS J081617.73+293639.6	0.768	1162	9.77	0.319
SDSS J133516.17+183341.4	1.192	1724	9.76	0.263

Table 1

G15 candidates providing the largest contribution to the expected GWB. Columns give the object identification name, redshift, orbital period in days, total mass as given in G15 and individual contribution to the GWB $A_s(f = 1 \text{ yr}^{-1})$ in Model 2b.

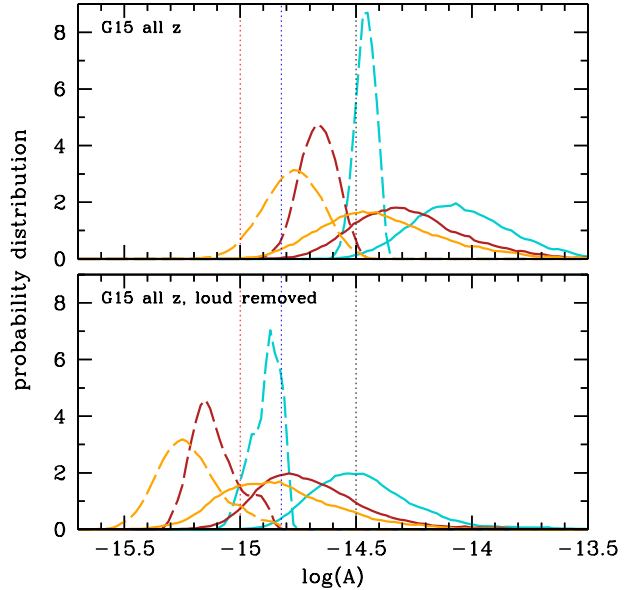


Figure 8. Same as Fig. 6, but investigating the weight of different candidate sub-populations. The upper panel shows the signal inferred from all sources, including those at $z > 1.3$; the lower panel addresses the effect of removing the 11 systems which contribute most to the predicted GWB. Line style as in Fig. 6.

the differential density of mergers per unit redshift and chirp mass, $d^2n/(dzd\mathcal{M})$. Each individual candidate in the CRTS sample contributes to this quantity via equations (2) and (9). The latter includes the dz/dV_c factor accounting for the co-moving volume shell accessible at a given redshift: a single binary observed at higher redshift implies a lower merger rate density because of the larger comoving volume available. One might therefore think that higher z systems do not contribute significantly to the GW signal. Since for a fixed observed frequency binaries at higher redshift are emitting at higher rest-frame frequencies, however, their intrinsic coalescence time is shorter, and the inferred merger rate larger according to equation (2). It turns out that this latter fact compensates for averaging over a larger volume shell, and the GW signal is dominated by SMBHBs at $z > 1.3$.

To demonstrate this, we performed the following experiment. Under the assumption of Model 1b, we generated 1000 realisations of M_1 and M_2 of each of the 98 CRTS binaries. We then used equations (2) and (9) to compute their contribution to the SMBHB merger density, and folded the result into equation (7) to compute the associated GW signal. We stress again that although the merger rate obtained via equation (2) depends on the chosen bin size, this is compensated by the integral over the size of the bin in equation (7). For each individual binary we compute $A_{50\%}$ (i.e. the median h_c at $f = 1\text{yr}^{-1}$; cf. equation (10) over the 1000 realisations. The distribution of those medians is shown in the upper panel of Fig. 7. The separation of the $z < 1.3$ and $z > 1.3$ populations reveals that members of the latter generally contribute the most to the GW signal.

Another interesting fact, highlighted also by the cumulative amplitude distribution of the individual SMBHBs, is that there appears to be a long tail at $A > 3 \times 10^{-16}$, implying that a handful of systems (11, for the record) contribute the vast majority of the inferred GWB. Therefore, in principle, misclassification of a relatively small number of candidates might have a strong impact on the results. These systems are listed in Table 1 and not surprisingly are among the most massive candidates. High masses imply short merger time-scales ($< 100\text{yr}$ for some of them assuming $q = 1/2$, cf G15) and consequently a high contribution to the merger rate. Interestingly, six of these sources are in a fairly narrow, high redshift range $1.7 < z < 2.3$. As we show below, these systems contribute about 2/3 of the overall GWB.

We explore the effect of including $z > 1.3$ sources and excluding the tail of loudest systems separately in Fig. 8. The top panel of the figure highlights the unexpected importance of $z > 1.3$ sources. When included, the signal is boosted up by almost 0.3dex, i.e. a factor of two. This is in sharp contrast with the results of the aforementioned theoretical models of SMBHB evolution. When high redshift sources are included, the median GWB amplitude values range from $\log A = -14.75$ for Model 3a to $\log A = -14.1$ for Model 1b, always much larger than the best PTA upper limits. When the 11 loudest sources are (arbitrarily) removed, however, the signal is reduced by almost 0.5dex, i.e. by a factor of three. The median values of the signal now range from $\log A = -15.2$ for Model 3a to $\log A = -14.55$ for Model 1b. We notice that most of the distributions are still inconsistent with the PPTA upper limit at $\log A = -15$, especially when mass measurement uncertainties are included (i.e. ‘b’ models).

The above exercise shows that misidentification of a few specific systems might have a significant impact on the GW predictions. It is therefore interesting to investigate how individual sources contribute to building up the signal. We turn to the more conservative Model 2b, which has a q distribution consistent with cosmological simulations of galaxy evolution (therefore less biased towards equal mass binaries than Model 1b). We rank the sources by individual contribution to the GWB and compute 1000 realization of the signal progressively removing sources one by one from the loudest to the quietest. The result is shown in Fig. 9 for the progressive removal of the loudest 50 systems. The overall signal drops steeply as the first 11 sources are removed. Their median contribution to the GWB for this specific model is listed in the last column of Table 1. Interestingly, SDSS J164452.71+430752.2 alone implies a median GWB of the order of the tightest PTA upper limit. Although these 11 sources contribute the majority of the signal, the GWB is still well

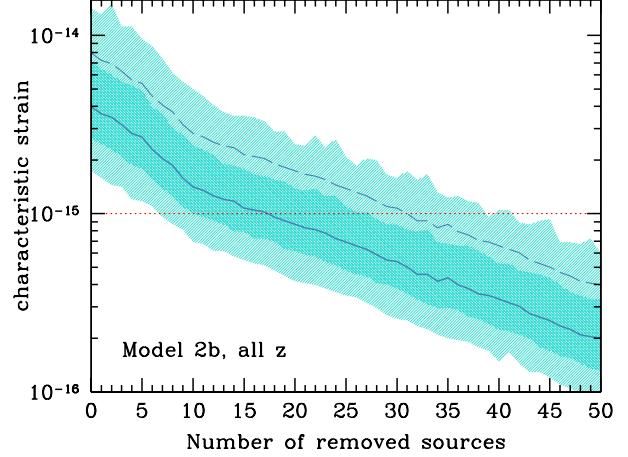


Figure 9. GWB build-up. Plotted is the characteristic strain of the GWB at $f = 1\text{yr}^{-1}$, A , versus the number of removed CRTS candidates. Candidates are ranked and removed, starting with the loudest, from left to right. Shaded areas mark the 68% and 95% confidence region of the amplitude computed over 1000 realizations of the signal, and the solid blue line is the median value. The dashed blue line is the median shifted upwards by a factor of two to account for incompleteness (See discussion in Section 4.3). The horizontal dotted red line is the PPTA upper limit. Model 2b is assumed.

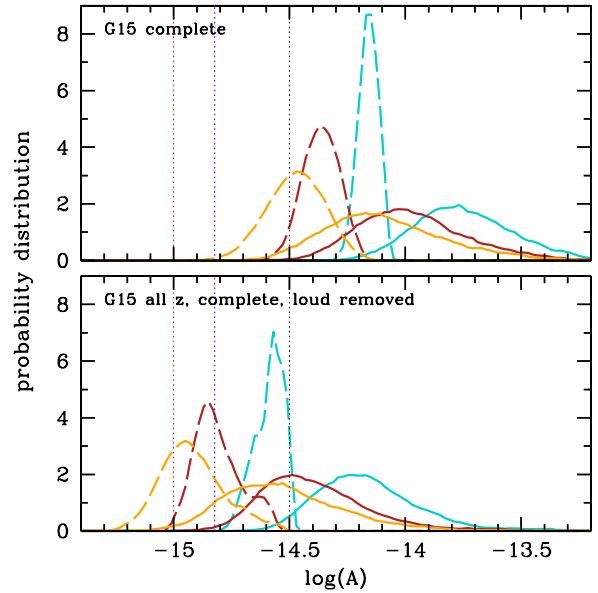


Figure 10. Same as Fig. 8, but the signal has now been corrected upwards by a factor of two to account for the approximate 25% completeness of the CRTS sample. Line style as in Fig. 6.

above current limits when they are removed. Moreover, we will see below that the CRTS sample is severely incomplete in many respects, and the implied GWB is *at least* twice as large. The dashed line in Fig. 9 shows the median of the signal when a factor of two correction is applied. Only a misidentification of the loudest 30+ SMBHBs would make the median GWB implied by the CRTS sample consistent with current PTA upper limits. This is strong evidence that most of the candidates cannot be SMBHBs and their observed variability must have a different origin.

4.3. Correcting for incompleteness

While these results are already inconsistent with current PTA upper limits, the discrepancy is significantly worsened by the fact that the SMBHB sample is necessarily incom-

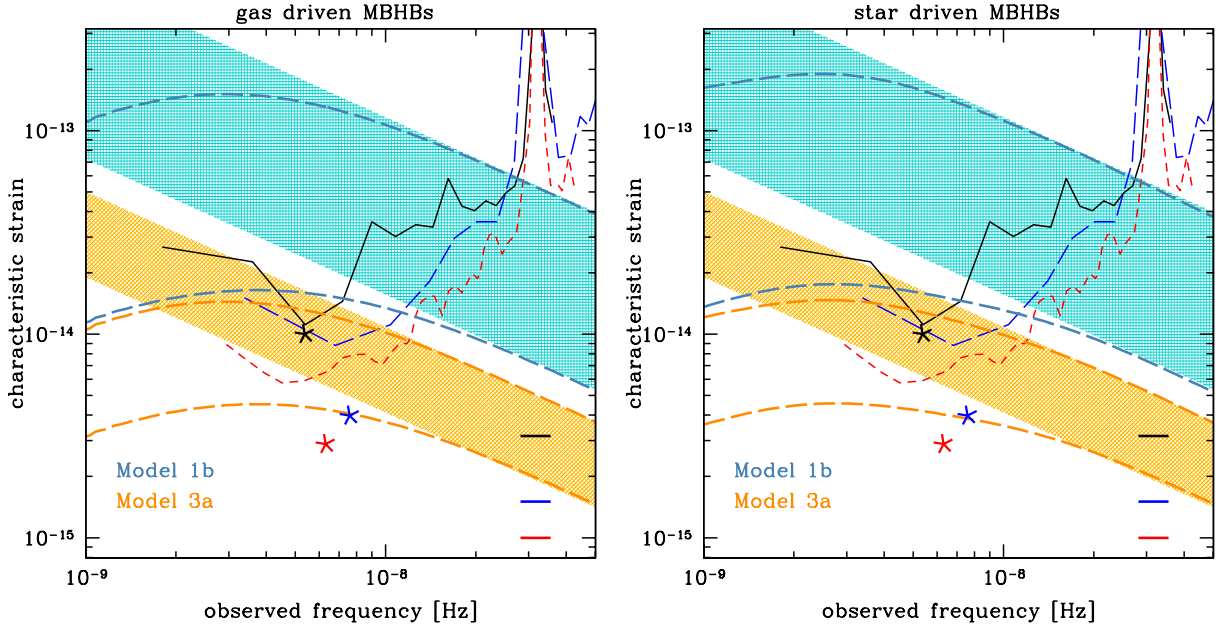


Figure 11. Effect of gas (left) and stellar (right) dynamics on the GW signal implied by the CRTS SMBHB sample (see text for details of the models of SMBHB-environment couplings). In each panel, the blue and orange shaded areas are the 95% confidence intervals produced by Model 1b and Model 3a respectively. Signals were computed including G15 SMBHBs at all redshifts and correcting for incompleteness of the sample (see text for full details). The long-dashed blue and orange lines represent the 95% predicted range for h_c under the effect of environmental coupling. The representation of the PTA sensitivity curves and upper limits is the same as in Fig. 5.

plete in several ways. First, CRTS covers only 80% of the sky (about 33000deg^2). Although this is a significant fraction of the whole sky, most of the selected quasars are identified in the SDSS survey (78 out of 111), which covers only about one quarter of the sky. Assuming a complete identification in the SDSS field of view, we would therefore expect $78 \times 4 = 312$ objects in the whole sky. We conclude that the G15 catalogue is at best only 35.5% complete based on sky coverage arguments only. Moreover, of the 334k identified quasars, 83k (25%) were rejected because of sparse sampling. Finally, we dropped the 13 candidates with no mass estimate (12%) of the sample. Adding everything up, the sample we use is *at best* only 24% complete. Since the signal is proportional to the square root of the coalescence rate, all the signal estimates shown so far must be multiplied by *at least* a factor of two.

The effect of incompleteness is shown in Fig. 10. The GWB predicted from the G15 sample is now inconsistent by a factor that ranges from 3 to 15 depending on the model. Even an ad hoc removal of the 11 systems accounting for most of the signal leaves us with a severe signal overestimation, with only Model 3a marginally consistent with the PPTA upper limit.

Besides these simple ‘counting arguments’ which can be easily corrected for, there are potentially many more sources of incompleteness in the G15 SMBHB sample. First, the redshift distribution of the CRTS candidate quasars shows a prominent peak around $z \approx 0.8$, hinting to incompleteness at higher redshifts. Second, the candidates are all spectroscopically confirmed type 1 quasars; numbers should therefore be corrected for the fraction of obscured type 2 quasars, which is poorly known but can easily double the sample (Lusso et al. 2013). Third, the search is constructed to look for sinusoidal periodicities and would be much less sensitive to eccentric binaries. Finally, not all SMBHBs have to be active in the first place, especially in gas poor mergers that are rather frequent at low redshift. In any case, the factor of two-corrected signal shown in Fig. 10 is necessarily a lower limit to the intrinsic

GWB implied by the CRTS SMBHB candidate population.

4.4. Possible impact of SMBHB coupling with stars and gas

In the previous discussion, we considered circular GW-driven binaries. However, it has been shown that both coupling with the environment and large eccentricities can in principle suppress the GW signal at nHz frequencies (see e.g. Enoki & Nagashima 2007; Kocsis & Sesana 2011; Sesana 2013a; Ravi et al. 2014).

Since the G15 SMBHB candidates are quasars, it is important to explore the case of strong coupling with a gaseous environment. Relevant models were constructed by Kocsis & Sesana (2011). They studied the SMBHB-disc coupling assuming different thin disc models (Shakura & Sunyaev 1973) described in Haiman et al. (2009). In particular they considered steady-state solutions by Syer & Clarke (1995), assuming both α and β discs, and the self-consistent non-steady solution of Ivanov et al. (1999). They found that only α discs with a large viscosity parameter ($\alpha_{SS} = 0.3$, in that work) and large accretion rate ($\dot{M}/\dot{M}_{\text{Edd}} = 0.3$ in that work) can significantly suppress the signal at the frequencies where current PTAs are sensitive. We therefore consider this model here, even though it might be unlikely to occur in nature, because it is not clear whether α discs can be secularly thermally stable (see, e.g., Hirose et al. 2009; Jiang et al. 2013). Moreover, we add eccentricity to the binary, according to findings of Roedig et al. (2011); binaries have $e \approx 0.6$ as long as they are coupled with the disc, and they progressively circularize because of GW emission after decoupling. Results are shown in the left panel of Fig. 11 for both Model 1b and Model 3a (i.e. the two models bracketing the expected GWB amplitude). Even considering such an extreme coupling, the implied GW signal is reduced by a factor of ≈ 2 at 6 nHz (and ≈ 7 at 1 nHz) and is still severely inconsistent with PTA upper limits. We stress that any other model (i.e., alpha discs with a smaller α -viscosity, any β disc, or any non-steady self-consistent disc)

would result in a negligible suppression of the signal in the PTA frequency range.

Coupling with stars can also accelerate SMBHB evolution and promote eccentricity growth (see, e.g., Quinlan 1996; Sesana 2010), both of which suppress the low frequency GW signal. Sesana (2010) modelled the stellar environment as a broken power-law, with a central cusp with density $\rho \propto r^{-\gamma}$ joining an external isothermal sphere at the SMBHB influence radius. In those models the SMBHB loss cone is assumed to be always full, which implies that the binary evolves at a pace that is dictated by the stellar density ρ_0 at the influence radius. Since the isothermal sphere is quite compact, especially compared to observed inner density profiles of massive ellipticals, these models can overestimate ρ_0 by more than an order of magnitude, and are therefore optimistic in terms of efficiency of the SMBHB-stellar coupling. We consider here the model with $\rho \propto r^{-1}$ and $e_0 = 0.6$ (e_0 is the eccentricity at binary formation). Results are shown in the right panel of Fig. 11 again for both Model 1b and Model 3a. Also, the suppression of the GWB in this case is milder, and both models are still inconsistent with PTA upper limits.

In principle, a larger suppression is possible if SMBHBs get more eccentric. However, the Bayesian analysis performed by D’Orazio et al. (2015b) on PG1302 showed that its eccentricity is at most ≈ 0.2 and it is consistent with zero. All SMBHBs in CRTS have been identified by their sinusoidal behaviour, which implies small eccentricities. The models considered in this section are consistent with SMBHBs having $e < 0.25$ at $f > 10^{-8}$ Hz, i.e. in the frequency range probed by CRTS. For example, a stellar driven model with $e_0 = 0.8$ would result in a population of fairly eccentric binaries at $f > 10^{-8}$ Hz, inconsistent with the sinusoidal light curves observed in the CRTS sample.

We caution that our computation *underestimates* the signal in the case of extreme environmental couplings. In fact, the normalization of the signal is set by the merger rate that is computed according to equation (4) assuming circular GW-driven binaries. If SMBHBs are still coupled with the environment, then the coalescence time $T_c(f)$ would be shorter, resulting in an higher merger rate and, in turn, in an higher signal normalization.

5. APPLICATION TO THE PTF SAMPLE

We also apply the same methodology described in Section 3 to the PTF sample identified by C16. By studying a sample of 35,383 spectroscopically confirmed quasars from PTF, they construct a sample of 33 binary candidates with periods $\lesssim 500$ days. They find that the coalescence time-scale distribution of the candidates best matches the expectations from a population of SMBHBs if their mass ratio is $q \approx 0.01$. We therefore constructed mock merger rates assuming $\bar{P}_{\text{orb}} = 500$ days and $q = 0.01$ and generated 1000 realizations of the expected GWB. Note that, in terms of GW emission, this model is extremely conservative. When any of the q distribution previously investigated is applied, the resulting GWB is much higher. As shown in Fig. 12, taken at face value, this sample is broadly consistent with PTA upper limits, mostly by virtue of the small q . We notice, however, that this sample has a severe incompleteness. Of the 33 candidates, 28 fall in the SDSS footprint. Considering that the PTF and SDSS footprints overlap over 9700 deg^2 , one would expect $28 \times (9700/41253) = 119$ systems for a uniform sky coverage. Moreover, of the 278,740 spectroscopically confirmed quasars in PTF, only 35,383 passed the selection cri-

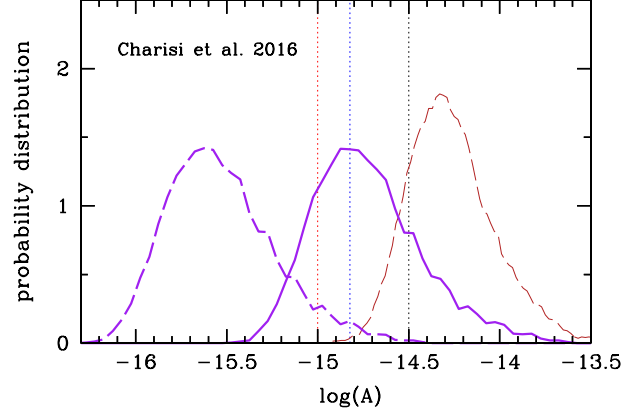


Figure 12. Distribution of the stochastic GWB characteristic amplitudes inferred from the PTF sample. The dashed purple curve is the GWB directly implied by the C16 sample, whereas the solid line accounts for incompleteness of the sample. The GWB produced by Model 2b built on the CRTS sample is shown as dashed brown line for comparison. The vertical dotted lines are 95% upper limits given by EPTA (black), NANOGrav (blue) and PPTA (red).

teria in terms of observational cadence. Considered together, these two facts imply that the C16 sample is only 3% complete. This implies a factor of $\sqrt{40} \approx 5.8$ underestimation of the GWB. When this correction is applied, the resulting amplitude distribution is largely inconsistent with the PPTA upper limit.

A few more things should be noted. On the positive side, being $q = 0.01$, these binaries are likely to be in the gas-driven rather than GW-driven hardening regime. This means that the spectrum might be much flatter than the canonical $f^{-2/3}$ power law. Moreover, those systems have GW frequencies $f > 4 \times 10^{-8}$ Hz, more than a factor of 5 larger than the frequencies at which PTAs are more sensitive (asterisks in Fig. 5), making extrapolation at low frequency critical (contrary to the CRTS sample). Extrapolating a flat spectrum instead of an $f^{-2/3}$ power by a factor of five in frequency would result in a constraint that is a factor of ≈ 2.5 less stringent. Although this would make the PTF sample consistent with PTA upper limits, we stress that we did not correct for detection biases due to inclination in the Doppler boosting interpretation of the periodicity (see D’Orazio et al. 2015b). Moreover, we did not consider any correction for missing SMBHBs with $q > 0.01$; if Doppler boosting manifests itself only at these small q , then there must be a large population of undetected SMBHBs with larger q contributing to the GWB, likely boosting it to unacceptable levels. Alternatively, one has to envisage a model whereby only $q \approx 0.01$ binaries form, which would be unexpected in the context of hierarchical build-up of SMBHBs.

6. CONCLUSIONS

We have shown that the list of supermassive black hole binary candidates identified by Graham et al. (2015) in the Catalina Real-time Transient Survey is inconsistent with current pulsar timing array limits on a stochastic gravitational wave background at nHz frequencies. Although none of the candidates, taken individually, is inconsistent with PTA measurements, they can be collectively used to construct a cosmic SMBHB merger rate and the associated GWB at nHz frequencies. The GWB computation implies the knowledge of the chirp masses of the candidates; we therefore need to assign a mass ratio to each object. We explore different possible

mass ratio distributions and we also tested models in which we allow a 0.3dex error on the mass measurements reported in G15.

Our main finding can be summarized as follows:

1. Even under the (unrealistic) assumption that the G15 sample is complete, and when only $z < 1.3$ sources are considered, the inferred GWB is hardly consistent with the most stringent PTA upper limit (*cf.* Fig. 6).
2. Contrary to theoretical expectations, SMBHBs at $z > 1.3$ dominate the GWB (*cf.* Fig. 8), increasing the total GWB by almost a factor of two, implying a factor of 2-to-8 inconsistency with PTA upper limits.
3. About 2/3 of the signal is contributed by a set of 11 SMBHB candidates (*cf.* Fig. 7 and Table 1). Misidentification of few particularly loud systems can therefore have a significant impact on the expected GWB.
4. When corrected for sky coverage and other selection effects, the GWB is increased by another factor of two. Considering all investigated models, the expected median amplitude of the GWB normalized at $f = 1\text{yr}^{-1}$ lies in the range $-14.45 < \log A < -13.8$. This is a factor of 3-to-15 higher than PTA upper limits, and still does not take into account that likely not all SMBHBs are Type-1 quasars.
5. Environmental coupling can decrease the signal by at most a factor of two at 6 nHz (i.e. where current PTAs are most sensitive), leaving therefore a severe discrepancy (at least a factor of 2-to-5) with PTA measurements. We stress that this estimate is optimistic, because it does not take into account that coupling with the environment would shorten the coalescence time scale of the SMBHB candidates, resulting in a higher merger rate and, in turn, in a higher GWB normalization (which we ignored in Section 4.4).

We stress that we implicitly assumed that *all* SMBHBs with orbital periods of few years are periodic quasars. If only a fraction of those SMBHBs exhibit AGN activity than the predicted stochastic GW background violates the current PTA upper limits even more.

We note that both the assumed q distribution and the inclusion of errors in the mass measurements have an important impact on the results. The latter tends to add uncertainty in the inferred rates, and allows for more massive binaries. In fact, ‘b’ models have a broader A distribution, with a median value shifted by about +0.3dex (i.e. a factor of two higher). The range of q distributions explored here further affect the expected GWB amplitude by an additional 0.5dex (i.e. a factor of three).

These results show that the SMBHB hypothesis for the CRTS sample, as presented in G15, is not tenable. To get some insight on how this candidate sample can be reconciled with current PTA upper limits, we turn to equation (7). This equation shows that h_c is proportional to the square root of the SMBHB merger rate, which is in turn proportional to the number of SMBHB candidates as per equation (3). An average factor 3-to-15 suppression is therefore achievable if at least 90% of the systems are not SMBHBs, which leaves us with at most a handful of them (in the best case). Note, however, that the contribution to the GWB varies significantly

across candidates (see Fig. 7). If the ≈ 30 loudest systems are excluded, the GWB implied by the leftover ≈ 70 would be consistent with current PTA upper limits even after incompleteness correction (*cf.* Fig. 9).

If we want to leave the number of SMBHB candidates untouched, the other parameter to look into is the inferred mass of each system. Equation (7) implies in this case $h_c \propto \mathcal{M}^{5/3}$. A factor $\mathcal{M}^{5/6}$ comes from the square root of $dE/d\ln f$ and another $\mathcal{M}^{5/6}$ contribution comes from equations (3) and (2); if the candidate masses are lower, the coalescence time scale is longer and the implied merger rate smaller. Suppressing the resulting GWB by a factor 3-to-15, would therefore require an overestimation of all candidate masses of at least a factor that ranges between 2 and 5 (always > 4 when mass measurement errors are taken into account, i.e. when ‘b’ models are considered). Alternatively, the mass ratio distribution of all these systems have to be extreme. After correcting for incompleteness, we find that the G15 sample is 95% consistent with the 95% PPTA upper limit only if $q < 0.003$. A scenario in which all SMBHBs have such a small mass ratio is difficult to accommodate in current galaxy formation models. Moreover, it is not clear whether such small mass ratio perturbors would result in a luminosity modulation of about 50%, as observed in the CRTS sample.

Alternatively, it has been proposed that the optical variability of SMBHB may be related to the appearance of an $m = 1$ mode at the inner edge of the circumbinary disk (D’Orazio et al. 2015a; Miranda et al. 2017b). In this interpretation, however, the binary orbital frequency would be ≈ 5 times higher than the observed variability in the CRTS sample, implying typical coalescence rates a factor $5^{8/3} \approx 70$ larger and an associated GWB louder by almost an order of magnitude. It is clear that this interpretation of the CRTS candidates is not viable. Conversely, C16 mentioned the possibility that the periodicity might be related to higher harmonics of the orbital periods. This would of course implying much longer periods and coalescence times, making PTA GWB upper limits less constraining. Although this possibility cannot be excluded, simulations of SMBHBs in circumbinary disks generally show prominent periodicities related to the binary orbital motion and/or to the circumbinary disk inner edge.

A similar analysis applied to the PTF sample identified by C16, yields comparable results. When corrected for incompleteness, the resulting GWB is a factor of ≈ 2 inconsistent with PTA upper limits. Although this might be mitigated by strong coupling with a circumbinary disk (suppressing the GWB at low frequency), our result was obtained under the assumption that variability is due to Doppler boosting and all candidates have $q = 0.01$. If Doppler boosting introduces a preferential bias towards small q , then the GWB has to be corrected for the missing fraction of SMBHBs with larger mass ratios. Alternatively, one has to put forward a model whereby only binaries with $q \approx 0.01$ form, which would be unexpected in the current framework of SMBH assembly.

In summary, we conclude that both SMBHB candidate samples are in severe tension with the current PTA measurements. The CRTS G15 sample implies a GWB that is a factor of 3-to-15 higher than the most stringent PTA upper limits, requiring that nearly all of the candidates are false positives. While this tension appears difficult to alleviate, possibilities include: (i) typical SMBH masses have been overestimated by a surprisingly large factor ($\gtrsim 4$), (ii) the typical mass ratio is surprisingly low ($q \lesssim 0.01$), or else that the loudest GW sources are preferentially false positives. The PTF C16 sample is already

a factor of ≈ 2 higher than the best PTA upper limit, even when assuming that all SMBHBs have $q = 0.01$.

While these results question the viability of SMBHB identification via periodicity studies alone, they demonstrate the status of PTAs as important astrophysical probes. Even without a direct GWB detection, PTA upper limits can put stringent constraints on interesting candidate objects. Further studies of these systems are required to identify the true origin of their periodic variability.

We thank Maria Charisi for useful discussions. AS is supported by a University Research Fellowship of the Royal Society. Financial support was provided by NASA through the ATP grant NNX15AB19G (ZH). ZH also gratefully acknowledges support from a Simons Fellowship in Theoretical Physics. This work was supported in part by the European Research Council under the European Union's Horizon 2020 Programme, ERC Starting Grant GalNUC #638435

REFERENCES

- Artymowicz, P., & Lubow, S. H. 1996, *ApJ*, **467**, L77
- Arzoumanian, Z., et al. 2014, *ApJ*, **794**, L41
- . 2015, *ApJ*, **813**, 65
- . 2016, *ApJ*, **821**, 13
- Babak, S., et al. 2016, *MNRAS*, **455**, 1665
- Barausse, E. 2012, *MNRAS*, **423**, 2533
- Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, *Nature*, **287**, 307
- Berczik, P., Merritt, D., Spurzem, R., & Bischof, H.-P. 2006, *ApJ*, **642**, L21
- Charisi, M., Bartos, I., Haiman, Z., et al. 2016, *MNRAS*, **463**, 2145
- Comerford, J. M., Schluns, K., Greene, J. E., & Cool, R. J. 2013, *ApJ*, **777**, 64
- Croton, D. J., Springel, V., White, S. D. M., et al. 2006, *MNRAS*, **365**, 11
- Cuadra, J., Armitage, P. J., Alexander, R. D., & Begelman, M. C. 2009, *MNRAS*, **393**, 1423
- Decarli, R., Dotti, M., Fumagalli, M., et al. 2013, *MNRAS*, **433**, 1492
- Desvignes, G., et al. 2016, *MNRAS*, **458**, 3341
- D’Orazio, D. J., Haiman, Z., Duffell, P., Farris, B. D., & MacFadyen, A. I. 2015a, *MNRAS*, **452**, 2540
- D’Orazio, D. J., Haiman, Z., & MacFadyen, A. 2013, *MNRAS*, **436**, 2997
- D’Orazio, D. J., Haiman, Z., & Schiminovich, D. 2015b, *Nature*, **525**, 351
- Dotti, M., Sesana, A., & Decarli, R. 2012, *Advances in Astronomy*, **2012**, 940568
- Drake, A. J., Djorgovski, S. G., Mahabal, A., et al. 2009, *ApJ*, **696**, 870
- Enoki, M., & Nagashima, M. 2007, *Progress of Theoretical Physics*, **117**, 241
- Eracleous, M., Boroson, T. A., Halpern, J. P., & Liu, J. 2012, *Astrophysical Journal Supplements*, **201**, 23
- Escala, A., Larson, R. B., Coppi, P. S., & Mardones, D. 2005, *The Astrophysical Journal*, **630**, 152
- Farris, B. D., Duffell, P., MacFadyen, A. I., & Haiman, Z. 2014, *ApJ*, **783**, 134
- . 2015, *MNRAS*, **446**, L36
- Foster, R. S., & Backer, D. C. 1990, *ApJ*, **361**, 300
- Goicovic, F. G., Sesana, A., Cuadra, J., & Stasyszyn, F. 2016, ArXiv e-prints, [arXiv:1602.01966 \[astro-ph.HE\]](https://arxiv.org/abs/1602.01966)
- Graham, M. J., Djorgovski, S. G., Stern, D., et al. 2015, *MNRAS*, **453**, 1562
- Gültekin, K., Richstone, D. O., Gebhardt, K., et al. 2009, *ApJ*, **698**, 198
- Guo, Q., White, S., Angulo, R. E., et al. 2013, *MNRAS*, **428**, 1351
- Haiman, Z., Kocsis, B., & Menou, K. 2009, *ApJ*, **700**, 1952
- Hayasaki, K. 2009, *PASJ*, **61**, 65
- Hayasaki, K., Mineshige, S., & Sudou, H. 2007, *PASJ*, **59**, 427
- Hennawi, J. F., Myers, A. D., Shen, Y., et al. 2010, *ApJ*, **719**, 1672
- Hirose, S., Blaes, O., & Krolik, J. H. 2009, *ApJ*, **704**, 781
- Ivanov, P. B., Papaloizou, J. C. B., & Polnarev, A. G. 1999, *MNRAS*, **307**, 79
- Jaffe, A. H., & Backer, D. C. 2003, *ApJ*, **583**, 616
- Jiang, Y.-F., Stone, J. M., & Davis, S. W. 2013, *ApJ*, **778**, 65
- Ju, W., Greene, J. E., Rafikov, R. R., Bickerton, S. J., & Badenes, C. 2013, *ApJ*, **777**, 44
- Kauffmann, G., & Haehnelt, M. 2000, *MNRAS*, **311**, 576
- Kelley, L. Z., Blecha, L., & Hernquist, L. 2017, *MNRAS*, **464**, 3131
- Kelly, B. C., Becker, A. C., Sobolewska, M., Siemiginowska, A., & Uttley, P. 2014, *ApJ*, **788**, 33
- Khan, F. M., Just, A., & Merritt, D. 2011, *ApJ*, **732**, 89
- Kocsis, B., & Sesana, A. 2011, *MNRAS*, **411**, 1467
- Komossa, S., & Zensus, J. A. 2016, in *IAU Symposium, Vol. 312, Star Clusters and Black Holes in Galaxies across Cosmic Time*, ed. Y. Meiron, S. Li, F.-K. Liu, & R. Spurzem, 13
- Kormendy, J., & Ho, L. C. 2013, *ARA&A*, **51**, 511
- Kozłowski, S. 2017, *A&A*, **597**, A128
- Lentati, L., et al. 2015, *MNRAS*, **453**, 2576
- Lu, K.-X., Li, Y.-R., Bi, S.-L., & Wang, J.-M. 2016, *MNRAS*, **459**, L124
- Lusso, E., Hennawi, J. F., Comastri, A., et al. 2013, *ApJ*, **777**, 86
- MacFadyen, A. I., & Milosavljević, M. 2008, *ApJ*, **672**, 83
- MacLeod, C. L., Ivezić, Ž., Kochanek, C. S., et al. 2010, *ApJ*, **721**, 1014
- Mayer, L. 2013, *Classical and Quantum Gravity*, **30**, 244008
- Miranda, R., Muñoz, D. J., & Lai, D. 2017a, *MNRAS*, **466**, 1170
- . 2017b, *MNRAS*, **466**, 1170
- Nixon, C. J., Cossins, P. J., King, A. R., & Pringle, J. E. 2011, *MNRAS*, **412**, 1591
- Noble, S. C., Mundim, B. C., Nakano, H., et al. 2012, *ApJ*, **755**, 51
- Phinney, E. S. 2001, ArXiv Astrophysics e-prints, [astro-ph/0108028](https://arxiv.org/abs/astro-ph/0108028)
- Planck Collaboration XIII. 2016, *A&A*, **594**, A13
- Preto, M., Berentzen, I., Berczik, P., & Spurzem, R. 2011, *ApJ*, **732**, L26
- Quinlan, G. D. 1996, *New A*, **1**, 35
- Rajagopal, M., & Romani, R. W. 1995, *ApJ*, **446**, 543
- Rau, A., et al. 2009, *Publications of the Astronomical Society of Pacific*, **121**, 1334
- Ravi, V., Wyithe, J. S. B., Shannon, R. M., & Hobbs, G. 2015, *MNRAS*, **447**, 2772
- Ravi, V., Wyithe, J. S. B., Shannon, R. M., Hobbs, G., & Manchester, R. N. 2014, *MNRAS*, **442**, 56
- Reardon, D. J., et al. 2016, *MNRAS*, **455**, 1751
- Roedig, C., Dotti, M., Sesana, A., Cuadra, J., & Colpi, M. 2011, *MNRAS*, **415**, 3033
- Roedig, C., Sesana, A., Dotti, M., et al. 2012, *A&A*, **545**, A127
- Runnoe, J. C., Eracleous, M., Pennell, A., et al. 2017, *MNRAS*, in press; e-print arXiv:1702.05465, [arXiv:1702.05465](https://arxiv.org/abs/1702.05465)
- Sesana, A. 2010, *ApJ*, **719**, 851
- . 2013a, *Classical and Quantum Gravity*, **30**, 224014
- . 2013b, *MNRAS*, **433**, L1
- Sesana, A., & Khan, F. M. 2015, *MNRAS*, **454**, L66
- Sesana, A., Roedig, C., Reynolds, M. T., & Dotti, M. 2012, *MNRAS*, **420**, 860
- Sesana, A., Shankar, F., Bernardi, M., & Sheth, R. K. 2016, *MNRAS*, **463**, L6
- Sesana, A., Vecchio, A., & Colacino, C. N. 2008, *MNRAS*, **390**, 192
- Sesana, A., Vecchio, A., & Volonteri, M. 2009, *MNRAS*, **394**, 2255
- Shakura, N. I., & Sunyaev, R. A. 1973, *A&A*, **24**, 337
- Shankar, F., Bernardi, M., Sheth, R. K., et al. 2016, *MNRAS*, **460**, 3119
- Shannon, R. M., et al. 2015, *Science*, **349**, 1522
- Shen, Y., Greene, J. E., Strauss, M. A., Richards, G. T., & Schneider, D. P. 2008, *ApJ*, **680**, 169
- Shen, Y., Liu, X., Loeb, A., & Tremaine, S. 2013, *ApJ*, **775**, 49
- Shi, J.-M., & Krolik, J. H. 2015, *ApJ*, **807**, L31
- Shi, J.-M., Krolik, J. H., Lubow, S. H., & Hawley, J. F. 2012, *ApJ*, **749**, 118
- Simon, J., & Burke-Spolaor, S. 2016, *ApJ*, **826**, 11
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Nature*, **435**, 629
- Syer, D., & Clarke, C. J. 1995, *MNRAS*, **277**, 758
- Tang, Y., MacFadyen, A. I., & Haiman, Z. 2017, *MNRAS*, submitted; e-print arxiv:1703.03913, [arXiv:1703.03913 \[astro-ph.SR\]](https://arxiv.org/abs/1703.03913)
- Tsalmantza, P., Decarli, R., Dotti, M., & Hogg, D. W. 2011, *ApJ*, **738**, 20
- Vasiliev, E., Antonini, F., & Merritt, D. 2015, *ApJ*, **810**, 49
- Vaughan, S., Uttley, P., Markowitz, A. G., et al. 2016, *MNRAS*, **461**, 3145
- Verbiest, J. P. W., et al. 2016, *MNRAS*, **458**, 1267
- Wang, L., Greene, J. E., Ju, W., et al. 2017, *ApJ*, **834**, 129
- White, S. D. M., & Rees, M. J. 1978, *MNRAS*, **183**, 341
- Wyithe, J. S. B., & Loeb, A. 2003, *ApJ*, **590**, 691
- Zhu, X.-J., Hobbs, G., Wen, L., et al. 2014, *MNRAS*, **444**, 3709